DOCUMENT RESUME

ED 271 672 CG 019 211

AUTHOR Munoz, Ricardo F.; And Others
TITLE Depression Prevention Research: Design, Implementation, and Analysis of Randomized Trials.
PUB DATE 24 Aug 85
NOTE 20p.; Papers presented at the Annual Convention of the American Psychological Association (93rd, Los Angeles, CA, August 23-27, 1985).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150) -- Collected Works - General (020)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Data Analysis; *Depression (Psychology); *Intervention; *Prevention; Research and Development; *Research Design; *Research Methodology

ABSTRACT

This document contains three papers concerned with prevention intervention research, a new area of depression research which has shown great promise for contributing new knowledge to the understanding of depression. The first paper, "Clinical Trials vs. Prevention Trials: Methodological Issues in Depression Research" (Ricardo F. Munoz), emphasizes that engaging in prevention/intervention research will allow perspectives which may not be evident using a treatment point of view. By actively seeking out high-risk groups, prevention/intervention research will contact a large proportion of untreated cases not represented in clinical research. The second paper, "Implementing a Randomized Prevention Trial in English, Spanish and Chinese" (Yu Wen Ying and Regina Armas), discusses the Depression Prevention Research Project, a project which involved English-, Chinese-, and Spanish-speaking adults not currently in psychotherapy. The methodology of the project is described in which subjects were screened for depression and were then randomly assigned to a control group or to either an intervention class or a videotape information presentation, both of which discussed self-control techniques effective in cognitive behavioral treatments of depression. The third paper, "Statistical Analysis Issues in Randomized Depression Prevention Trials" (Florentius Chan, et al.), discusses the issues of experimental design, outcome evaluation, and data analysis techniques in reference to the Depression Prevention Research Project. It is concluded that no experimental design is perfect for answering all questions and that the researcher's aims should be to develop the design that best isolates tests of major hypotheses, to measure outcome dimensions sensibly, and to use statistical analysis techniques that make real differences salient. (ABL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED271672

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

DEPRESSION PREVENTION RESEARCH:

DESIGN, IMPLEMENTATION, AND ANALYSIS OF RANDOMIZED TRIALS

Chair: **Manuel Barrera, Jr.**, Arizona State University

Presentations: **Ricardo F. Muñoz**, University of California, San Francisco
Clinical Trials vs. Prevention Trials:
Methodological Issues in Depression Research

Yu Wen Ying and **Regina Armas**, University of California, San Francisco
Implementing a Randomized Prevention Trial
in English, Spanish, and Chinese

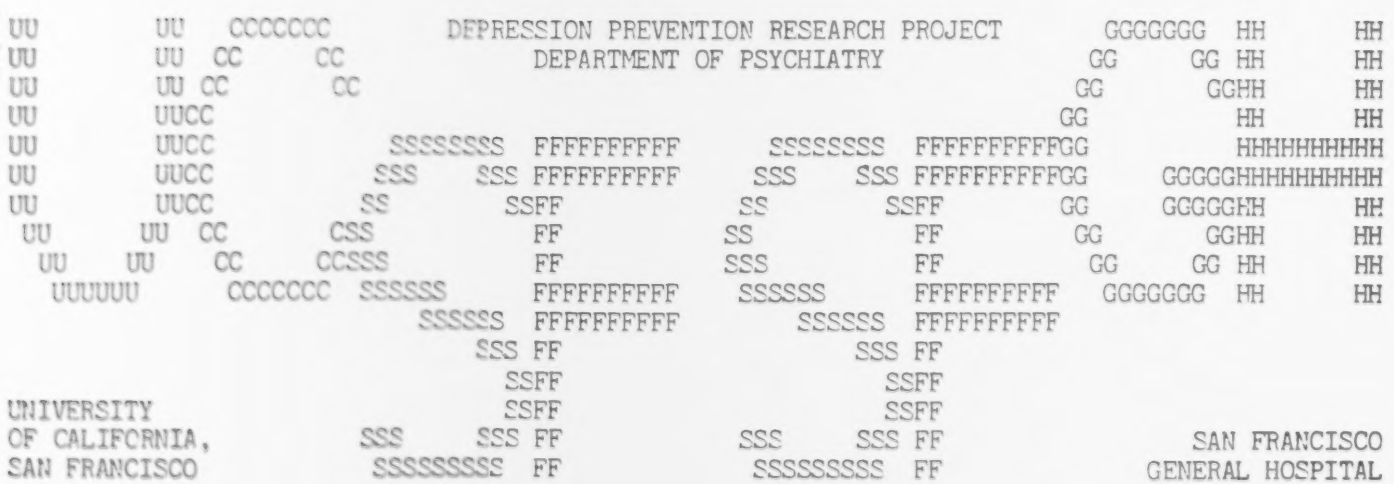
Florentius Chan, James L. Sorensen, and **William Hargreaves**,
University of California, San Francisco
Statistical Analysis Issues in Randomized Depression Prevention Trials

Discussant: **Richard H. Price**, University of Michigan

Symposium conducted at the meeting of the American Psychological Association,
Los Angeles, California, August 24, 1985

For further information, contact: Ricardo F. Muñoz, Ph.D.
University of California, San Francisco
Department of Psychiatry
San Francisco General Hospital (7G16)
San Francisco, CA 94110

CG019211


DEPRESSION PREVENTION RESEARCH PROJECT
DEPARTMENT OF PSYCHIATRY
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO
SAN FRANCISCO GENERAL HOSPITAL

Clinical Trials vs. Prevention Trials: Methodological Issues in Depression Research

Ricardo F. Muñoz, Ph.D.

University of California, San Francisco

Introduction to the symposium:

A number of significant advances have been made in the area of depression research in the past few decades. There is greater consensus regarding the diagnostic criteria for clinical depression. There have been a number of psychometrically acceptable self-report scales developed for both epidemiological and clinical studies. Pharmacological and psychological therapies have been found to be effective for unipolar depression in randomized clinical trials. In addition, biological, genetic, and social influences are being studied in progressively more sophisticated ways, and are yielding exciting information about depression.

In this symposium we will present the argument that a new area of research has been initiated which has great promise for contributing new knowledge to our understanding of depression. This area consists of prevention intervention research.

Our first paper will cover methodological issues in depression prevention trials, the second paper will discuss our experience in conducting such a trial, and the third will focus on the statistical analysis of prevention trials.

I will address the methodological issues by contrasting clinical trials (that is, treatment outcome studies) with prevention trials.

Representativeness. Investigators conducting clinical trials are more concerned with homogeneity in their sample than with the representativeness of the sample. The treatment outcome field has placed great emphasis on the use of structured interviews and diagnostic criteria which will ensure that all study participants suffer from clinical depression, usually defined as major depression.

This screening procedure produces samples of "pure depressives" defined by our present diagnostic categories. However, it does not produce representative samples of such persons from the community. For example, the Epidemiological Catchment Area Project reports that of those persons found to meet DSM-III criteria for a major depressive episode, only 20% visited a mental health professional. Out of this already small portion of depressed persons very few are likely to be informed about treatment studies taking place in their community, and even fewer are likely to make an inquiry about them. Of those who do make an inquiry, about 10% are scheduled for evaluation interviews (Bellack, Hersen, & Himmelhoch, 1981). Of those fully evaluated, approximately 36% are accepted into the randomized trial (Rush, Beck, Kovacs, & Hollon, 1977; Murphy, Simons, Wetzel, & Lustman, 1984). Further bias is introduced by attrition, which ranges between 20% and 52% in comparative treatment outcome studies of depression (Simons, Levine, Lustman, and Murphy, 1984; DiMascio, Weissman, Prusoff, Neu, Zwilling, &

Klerman, 1979). Thus, treatment outcome studies have populations which are unrepresentative of the general population of persons who meet major depression criteria.

Prevention intervention trials focus on populations at high risk for developing clinical depression. This has two implications: first, groups at risk are likely to include untreated individuals already meeting clinical criteria for depression. Therefore screening methods should be used to identify such persons and refer them for treatment. Second, because prevention trials are focused on demographically defined populations, they should attempt to obtain representative samples of these populations. Therefore, prevention trials must seek out persons at risk, rather that wait for them to contact the project.

The natural history of depression. Because depression is generally a time-limited disorder, and because individuals are most likely to come for treatment when they are feeling at their worst, the normal process of the disorder works in conjunction with the treatment. The episode is generally reaching its peak when persons come for treatment, and it is likely to diminish eventually, even if nothing is done. This is one of the reasons that controlled studies are so important. Successful treatments generally shorten the depressive episode, and thus differences between treated subjects and control subjects in the rate of improvement become evidence for the effectiveness of the treatment.

In prevention trials, specially ones in which persons

already depressed have been referred elsewhere for treatment, the participants are not in the middle of a depressive episode, but merely at risk. There are likely to be two subgroups: Those who are very unlikely to become depressed in the foreseeable future, and those who are very likely to become depressed, but have not become depressed yet. Those in the first group (which is likely to be the majority) will show little change throughout the trial (in either the experimental or the control condition), and thus will reduce the differential effect size. Those who are very likely to become depressed will be encountered prior to reaching their depressive peak, and thus the intervention must work against the natural course of the disorder. The intervention is therefore effective if the number of people who become depressed is less in the experimental condition than in the control condition.

To show such an effect, the high risk group must be carefully chosen, so that a large enough proportion of the group is likely to become depressed during the period which the trial is taking place. Therefore, preventive trials can allow us

1) to study the course of depressive disorders throughout their cycle, and not just from the time that the full blown disorder is identified, and

2) to study the differences between those individuals in high risk groups who become depressed versus those who do not.

Note that regression to the mean works in favor of treatment trials: since all the people have high levels of depression at the start, it is likely that their mean as a group will be closer to the normal population mean later. Effective treatments

accelerate that trend. Regression to the mean works against prevention trials. Since the extremely high depressive scores have been screened out and referred for treatment, we must await changes away from the mean in the control group to show a preventive effect in the experimental group.

The use of continuous versus dichotomous variables also produces differences in the evaluation of treatment and prevention studies. Most randomized depression treatment studies to date have consisted of under two hundred patients, and rarely more than fifty patients per treatment group. The dependent variable used to determine effect is a score on symptom scales. The reason that prevention trials need such large numbers (and we will refer to our own power analyses in the third paper), is that we set up reduction in incidence as the criteria for success. Yet we must realize that current diagnostic categories behave as continuous scales. In DSM-III, for example, the difference between a case and a non-case is merely the absence of one symptom out of many.

Conclusions. It is our contention that by engaging in prevention intervention research, we will force ourselves to take perspectives which may not be evident when approaching depresssion from a treatment point of view. These new perspectives may help to improve the methodological aspects of both clinical and prevention endeavors.

Prevention intervention research can be conceptualized as focusing on an area between epidemiology and clinical research. It is concerned with populations and representativeness, but

because it evaluates interventions, it must rely on samples of volunteers. Yet, by actively seeking out high risk groups, it will come into contact with the large proportion of untreated cases which are not represented in clinical research.

At some point, a merging of the two perspectives may be most profitable. Imagine, for example, a prevention study in which identified untreated cases are referred for treatment, but allowed to participate in the prevention study. Part of the evaluation might consist of the number who actually followed through and obtained clinical treatment in the experimental and control groups, the number who improved in each group with or without therapy, as well as the number of new cases. Such a design would test the effect of the intervention at a number of levels, such as compliance with referral, the "therapeutic side effects" of prevention interventions, and the preventive effects themselves.

We hope that other segments of the depression research field will welcome prevention intervention researchers into their ranks, and that the potential contributions of this new area will be recognized and encouraged.

Implementing a Randomized Prevention Trial in English, Spanish and Chinese

Yu Wen Ying, Ph.D.

Regina Armas, Ph.D.

University of California, San Francisco

In this presentation, we will briefly discuss the implementation of the Depression Prevention Research Project and major issues which emerged from that process. With the exception of the Chinese sample all subjects were recruited from outpatient medical clinics. The inclusionary criteria were 1) ages between 18 and 69, 2) not currently in psychotherapy, and 3) a medical chart open for at least six months for the English and Spanish speaking samples.

Recruitment Three methods of recruitment were used: Face-to-face, mailing and mass media. For the English- and Spanish-speaking samples, face-to-face and mailing recruitment were used. In face-to-face recruitment, medical outpatients were approached in the waiting room of the clinic. In mailing recruitment, names and addresses were obtained from the hospital billing department. Patients were then sent information about the study. Because of the limited number of Chinese speaking medical outpatients at the medical clinics where we recruited the English- and Spanish-speaking subjects, mass media was utilized to recruit this group. Advertisements about the study were placed in Chinese language newspapers. Interested readers were invited to contact the project and participate in the screening process.

Screenings Three screening sessions were conducted. At the first screening, demographic information was collected and the Center for Epidemiological Studies - Depression Scale (Redloff 1977) was administered. The second screening session consisted of the Beck

Depression Inventory (Beck, et al, 1961) and the NIMH Diagnostic Interview Schedule (Robins, Helzer, Ratcliff, 1981). The latter yielded DSM-III diagnoses on major disorders, including major depression and dysthymia. The third screening session consisted of measures assessing variables that are theoretically related to depression.

Participants were paid ten dollars for each of the two latter screenings. Persons meeting criteria for a current (i.e., within the last six months) diagnosis of depression (major depression or dysthymia) and mania, were excluded and referred for treatment to ensure the preventive nature of the study.

Randomization and Intervention At the conclusion of the screening, participants meeting exclusionary criteria were offered the opportunity to participate in the randomization. Subjects were randomized into the class or control conditions. With the English-speaking group, there were two control conditions: no intervention or an information-only condition, which consisted of a 40 minute videotape presentation of the main techniques discussed in the class intervention. The class intervention was a two-hour eight week course which discussed self-control techniques found to be effective in cognitive behavioral treatments of depression. The course was taught in groups of ten or fewer students.

Follow-up Follow-up interviews were administered immediately after the intervention and at six and twelve months from the time of the original DIS. Subjects were paid ten dollars for each of the follow-up interviews.

Recruitment Mailing was relatively inexpensive when compared to the recruitment hours needed for face to face recruitment. It has the advantage of requiring an independent and active participant role on the part of the outpatient from the very start (i.e., mailing

back the first screening materials). Face to face recruitment, on the other hand relies more heavily on the skill of the recruiter to elicit interest in the project. The personal touch of the recruiter may persuade participation from an otherwise not so interested outpatient. This may result in the subject not completing the screening once the contact with the recruiter is terminated. As we had suspected mailing recruitment yielded more conscientious subjects with a higher completion rate of screening than the face to face recruited subjects for this very reason (i.e., 57% vs. 47%).

Subject Retention Once a subject was recruited, a great deal of effort was spent to maximize retention. On the whole, this is a group that is likely to have had little prior contact with psychological research and for whom neither research nor mental health intervention were of high priority. Thus, coupled with severe stressors such as poverty, attendance at interview sessions was sporadic. In order to increase the attendance, we sent reminder letters and called subjects shortly before the appointment to remind them of the appointment and to encourage attendance. In the implementation of the classes, we sent weekly reminder letters regardless of subject's attendance. In addition, we called those who had missed a class. All these contacts proved to be helpful but not as successful as we would wish. For instance, of the subjects assigned to the class intervention 20% never came to class and another 20% came to fewer than four of the eight classes. One major problem is that with this population the primary concern may be economic survival and not psychological well-being. Future studies with this population might wish to collaborate with a setting that is already addressing some of these economic needs. For instance, in working with recent immigrants it would be advantageous to implement such a study in conjunction with an English as a Second Language program or a job training program. With a medical population one might present the intervention on a TV monitor in the waiting room. The key is to bring the intervention to the population and not to require special efforts on the part of the participants. In prevention where

the occurrence of the problem is only a possibility and not a certainty, it is crucial that participation require as little effort as possible.

For the subjects who did come, several reasons were cited for their participation. One was payment for interview sessions, another was the wish to contribute to the understanding of depression. Others noted that they had experienced depression in the past and wanted to learn how to prevent it from reccurring. This seemed to be the most highly motivated group since there was a match between the purpose of the research project and personal life experiences.

Working with Ethnically Diverse Populations The inclusion of diverse ethnic groups and non-English speaking samples raised the question of comparability of research instruments and intervention. Close attention was paid to the development of culturally appropriate translation of all research and intervention materials. For example, we used a strict forward and backward translation technique on all the materials. In addition, the staff who administered the measures and the intervention (especially for the non-English speaking samples) was chosen on the basis of not only fluency in the language but also personal experience with the subjects' cultures.

In delivering the intervention across the language groups, we found that certain modifications were necessary for the Spanish and Chinese-speaking populations. American culture focuses on the importance of the individual whereas Latin and Chinese cultures emphasize the importance of the family and role of the individual within the family system. Thus, instead of talking about what one can do to improve one's life and mood, we talked about this in the context of how these improvements would benefit a larger unit, i.e., the family, as well.

Certain cognitive behavioral concepts were found to be more culture-syntonic than others. For instance, with the Chinese-speaking group, increasing positive thoughts fits squarely with the emphasis on the utility of thinking and then behaving accordingly.

Assertion, on the other hand, is considered impolite and needs to be discussed in terms of the context where it might be appropriate and comfortable for the participant. With the Spanish-speaking group, the term "self-control" was found to be too constraining, and we chose instead to emphasize these methods as helping to increase one's freedom of action.

Conclusion In summary, we have considered three major questions in the implementation of a project such as ours. First, it is important to ask how one can best access the target population. Secondly, what are the major needs of the population and how does the intervention offered meet these needs? If it does not address the most urgent needs directly, can it be made part of a program which does? We talked about this in terms of how it affects subject retention, but clearly these questions are important in and of themselves. Finally, does the intervention match the population? What modifications may be needed to bring about a better match? Attention to these issues may not ensure a smooth implementation process but increases the probability of one.

References

Beck, A.T., Ward, C.H., Mendelson, M., Mock, J.E. and Erbaugh, J.K. (1961). Archives of General Psychiatry, 4, 561-571.

Radloff, L.S. (1977) The CESD-Scale: A self-report depression scale for research in the general population. Applied Psychological Measurement 1, 385-401.

Robins, L.N., Helzer, J.E., Croughan, J., and Ratcliff, K.S. (1981), National Institute of Mental Health Diagnostic Interview Schedule: Its history, characteristics and validity. Archives of General Psychiatry, 38, 381-389.

Statistical Analysis Issues in Randomized Depression Prevention Trials

Florentius Chan, Ph.D.
James L. Sorensen, Ph.D.
William A. Hargreaves, Ph.D.

This paper summarizes several major statistical concerns in studying prevention effects. Specifically, issues related to the experimental design, outcome evaluation, and data analysis techniques will be discussed with reference to the Depression Prevention Research Project.

Experimental Design

In this section, we will talk about the issues related to between-subject design, covariance, and repeated measures.

In prevention studies we often expect to see effects of small size or effects that do not appear for a long time. Both make it essential to preserve statistical power. One way to do this is to use the simple 2-group design, intervention vs. no intervention. Although this two-group design is simplest, it can be useful to add other levels to further refine the intervention effects. For example, the design of our Project is as follows: Half of the participants are randomly assigned to the class condition, and half to one of two control groups. The two control groups are evenly divided into a no intervention group and a minimal intervention group consisting of a 40 minute videotape which illustrates the ideas covered in the class. We conceptualize this as an information-only control group. By

using the ratio of 2:1:1, that is, two class subjects, one video subject and one control subject, we can partition the two degrees of freedom in the group comparison into "class vs. other" and "video vs. control". The logic of this study justifies treating these as planned comparisons, which means we would report one or the other of these tests as "significant" in the unlikely event that one of them was significant while the overall 3-group test was not significant. Both of these comparisons use the common error term from the ANCOVA, so subjects in all 3 groups contribute power to both tests. The "class vs. other" comparison retains the full statistical power of a 2-group design, if, in fact, video and control conditions have similar effects.

One of the major disadvantages of the between-subjects design is the problem of individual differences. Randomization helps to reduce the differences between groups but there still can be some chance baseline differences between groups. One solution is to control the effect of individual differences by adding covariates in the analysis. In our Project we measured several variables (e.g., social support, life events) which are believed to be related to depression. One or more of these variables could be used as covariates in the analysis to improve the sensitivity of the F test. However, by introducing covariates in the analysis, the researchers have to deal with the assumption of reliability of the covariates and homogeniety of covariance in addition to the homogeniety of variance. Obviously, there is no perfect solution and researchers have to

evaluate the pros and cons of the inclusion of covariates and the number of covariates to be included. The technique of "adaptive randomization" has recently gained popularity as a way to improve balance in randomized trials where stratification is impractical because subjects enter the study one at a time.

In prevention, the intervention is expected to reduce the probability of developing depression. The outcomes (e.g., reduction of incidence) are usually not salient immediately after intervention because the probability of occurrence of a depressive episode within a short period of time is relatively low. To detect differences between intervention and no intervention groups, long-term, repeated follow up assessments are necessary. The use of follow up assessments enables us to study one major within-subjects factor, that is, time. The time factor is particularly important in prevention studies. In our Project the 6-month and 1-yr follow up assessments were calculated from the intake assessments, which enabled us to compare incidence rates of depression to those found in epidemiology studies such as the Epidemiological Catchment Area Project. Since the intervention in our Project involves learning cognitive-behavioral techniques which are presumably new to the subjects, it is expected that the intervention effect would develop over a period of time. Thus there might be a sleeper effect. By using long-term, repeated follow up assessments, we would be able to pinpoint the time interval required to show the intervention effect. However, the risk in long-term follow up is high subject attrition.

Outcome Evaluation

There are three major outcomes in our Project, namely incidence of depression, depressive symptomatology, and intermediate variables.

The major purpose of prevention studies is to reduce the incidence of depression. Thus a major outcome measure is the incidence of depression. Although prevention studies usually involve high risk subjects, as in our Project, the incidence of depression is relatively low and a very large sample is needed to show an intervention effect on incidence of depression. For example, estimating a one-year incidence of 10% for depressive episodes, and a reduction of 50% in incidence (i.e., from 10% to 5% in the experimental group), our power analysis shows that to obtain power of .84, the total sample size needed would be 800. The NIMH reviewers have been unwilling to fund such a large study thus far.

One major problem in studying the episode of depression is that we categorize subjects into depressed vs. nondepressed based on the number or severity of depressive symptoms which are essentially continuous variables. However, by studying the continuous variables (mean depression scores) rather than the categorical variable (episodes of depression), the variability of the continous variables is taken into account and thus the sensitivity of the analysis is increased.

In addition to the depression levels, we are also interested in the intervention effect on health care costs with the assumption that depression inflates the health care costs.

Unlike the depression scales which are self-report measures, the health care costs based on the chart review provide an objective, indirect measure of distress. Health care costs also yield a quantitative scale that will allow good statistical power.

To better understand the mechanism of action of any treatment effect found, it is important to include process measures. Our intervention is based on the argument that changes on cognitive-behavioral variables would lead to the changes on depression variables. By studying the amount of variance in the changes on outcome variables accounted by the changes on process variables, we could assess their causal relationships.

Data Analysis Techniques

One of the greatest difficulties researchers encounter is the selection of the most powerful data analysis techniques. In our Project we are using correlational analysis for the preliminary examination of the relationships among independent variables and among dependent variables. Based on this information, we include or construct variables for the main tests of the hypotheses so that statistical power is maintained. For example, multivariate analysis of covariance can improve the sensitivity and power of the analysis if we know we have a cluster of related dependent variables. If we have several related independent variables of modest reliability, we may choose a structural equation approach to avoid spuriously large effects that can occur when one includes unreliable covariates. The general issue is to avoid using a large number of tests

because the likelihood of spurious "significance" increases with number of tests.

Conclusion

The presentation discussed three major statistical issues in conducting randomized depression prevention research, namely experimental design, outcome evaluation, and data analysis techniques. These issues are related to one another. For example, the selection of a data analysis technique is dependent on the experimental design, and the specific measurements of outcome dimensions. There is no experimental design or statistical technique that is perfect for answering all questions. The researcher's aim should be to develop the design that best isolates tests of major hypotheses, to measure outcome dimensions sensibly, and to use statistical analysis techniques that make real differences salient. If there is one piece of advice we can give, it is to be sure to include in your team from the initial planning stages a statistician well trained in the analysis of large field trials.